HDF-EOS Subsetting: HEW and other tools

HDF and HDF-EOS Workshop VI
San Francisco, CA – December 4-5, 2002

Dr. Sara Graves, Bruce Beaumont,
Helen Conover, Matt Smith
Presented by Steve Tanner
University of Alabama in Huntsville, Information Technology and Systems Center

SUBSET.ORG

Agenda

- Subsetting
- HDF-EOS Web-based subsetter (HEW)
- HEW Subsetting Appliance: Integration with ECS
- HSE: A new Subsetting Engine
- Software availability options
- HEW Demonstration

Subsetting

- **Goal: to provide a science data user with only the data they request as quickly as possible.**

- **Benefits science data users and data centers:**
  - **reduces analysis time by reducing amount of data**
  - **reduces time for data delivery**
  - **reduces resources (network, personnel, media, etc.)**

- **Steps:**
  - **locate spatial / temporal / spectral area of interest**
  - **extract**
  - **re-assemble for distribution/use**

Currently Available/Planned Subsetting Applications

- HDF-EOS Subsetting - HEW
  - Complete System (available)
  - Subsetting Back-end Only (available)
  - SPOT - Subsettability Checker (available)
  - Subsetting Center at UAH (available)
  - HEW Integration with ECS (in work)
  - HEW Subsetting Engine (in work)
  - Subsetting as a Web Service (planned)
- Customized Subsetting
  - MODIS tools (available)
  - Coarse-grain SSM/I Subsetter (available)
- General Purpose Customizable Subsetting
  - Based on ADaM Data Mining Engine (available)
  - Subsetting Tool using ESML (in work)

HEW

- HDF-EOS Web-based Subsetter
  - Dataset-independent for HDF-EOS files
  - Optional Front-end/GUI
    - Uses HTML forms and JavaScript
  - Back-end
    - Needs subset criteria and HDF-EOS data
    - Performs subsetting as a “batch” job
  - http://subset.org

HEW Back End

- Uses HDF-EOS (and HDF) library
- Instructed via a subset criteria file (ODL)
- Handles multiple similar files
- Handles Swath and/or Grid objects
- Unix (SGI & Sun) executables available
- Subsetted output files contain:
  - StructMetadata (HDF-EOS)
  - ArchiveMetadata*
  - ProductMetadata (added by HEW ← ODL file)
  - CoreMetadata* (w/ modified bounding box & time info)
    - optionally placed in `.met` file

** if present in parent file*

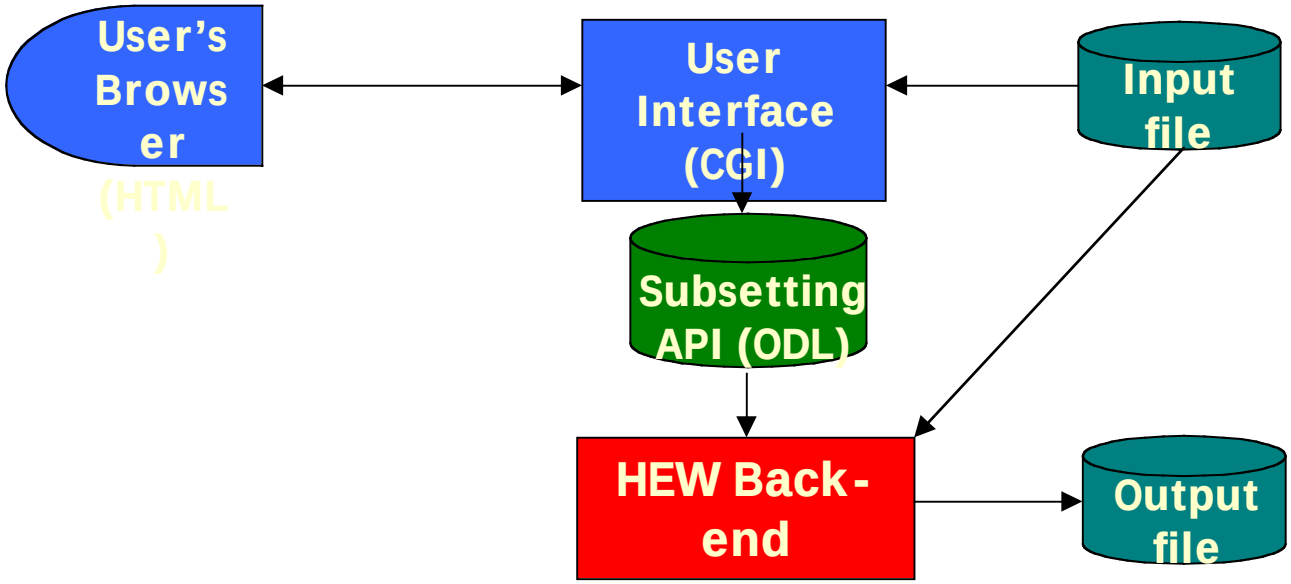
HEW Complete System

The User Interface checks the HDF-EOS file and presents the attributes to the user.

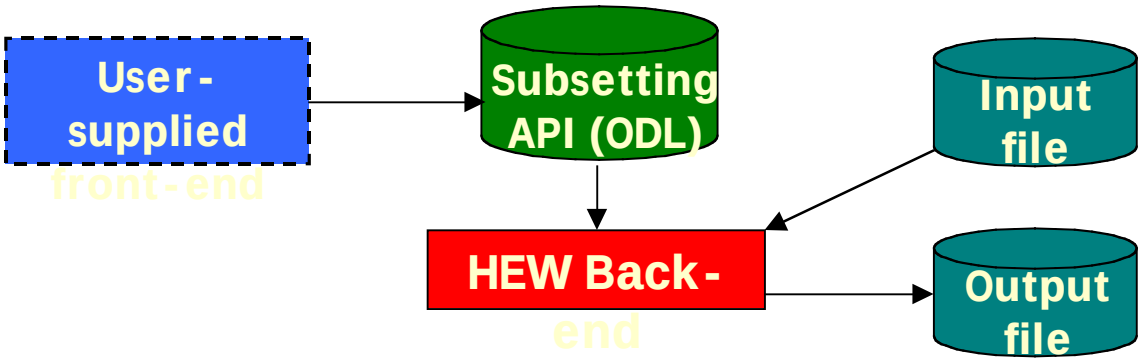
The user interacts with the browser to specify the subsetting criteria.

The User Interface creates the subsetting criteria file.

The HEW Back-end uses the subset criteria file with the input HDF-EOS file to create the subset HDF-EOS file.



HEW Back-end Only



A user-supplied program or other application creates the subsetting criteria file.

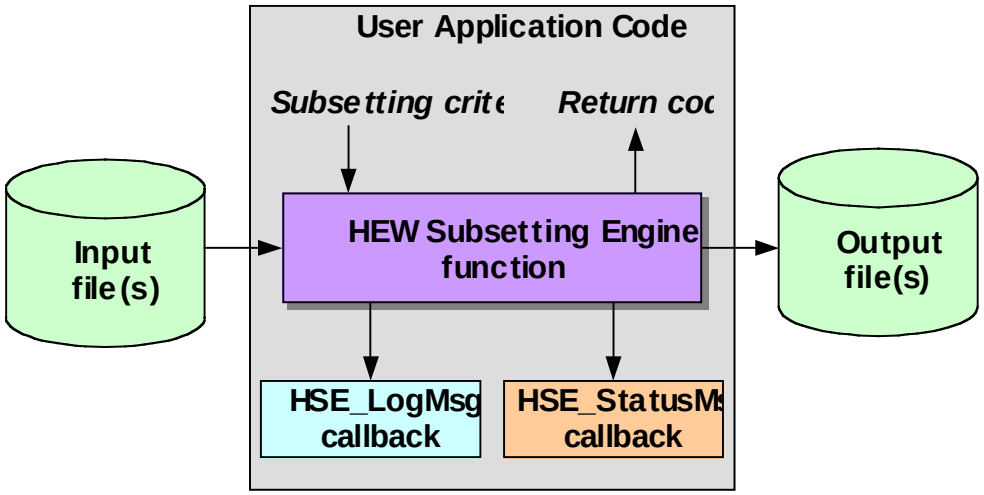
The HEW Back-end uses the subset criteria file with the input HDF-EOS file to create the subsetted HDF-EOS file.

Example Subset Criteria File

```
GROUP = SUBSET
        PARENT_FILE =("/AQUA/AMSR/AE_L2A.hdfeos")
        LATITUDE_RANGE = (35.000000, 40.000000)
        LONGITUDE_RANGE = (-77.000000, -72.000000)
        EMAIL = "matt.smith@msfc.nasa.gov"
        MET_FILE = YES
        GROUP = SPOG
                NAME = "swath_1"
                TYPE = "SWATH"
                PARAMETERS = ("89.0V_Res.1_TB",
                        "89.0V_Res.2_TB")
                SUBSAMPLING = ("GeoTrack", 2,
                        "GeoXtrack", 1)
        END_GROUP = SPOG
END_GROUP = SUBSET
END
```

HSE: HEW Subsetting Engine

Similar functionality to HEW back end, but contained within a callable *function*



- User application code builds subsetting criteria *structure*
- Subsetting engine *function* calls user's functions for status and log messages
- No e-mail is sent

- Beta version and documentation available now
  - Supports only grid subsetting at this time
  - BCEA and SOM grids cannot be subsetted yet
- Swath subsetting available first quarter 2003

HEW Subsettable data

EOS DATASETS

- **Terra**
  - **MODIS**
  - **MOPITT**
  - **ASTER**
- **Aqua**
  - **AMSR-E**
  - **AIRS**
- **Aura**
  - **HIRDLS**

OTHERS

- **TRMM**
  - **TMI**
- **NOAA-15,16**
  - **AMSU-A**

- any other HDF-EOS data written with HDF-EOS library subsetting calls in mind

SPOT

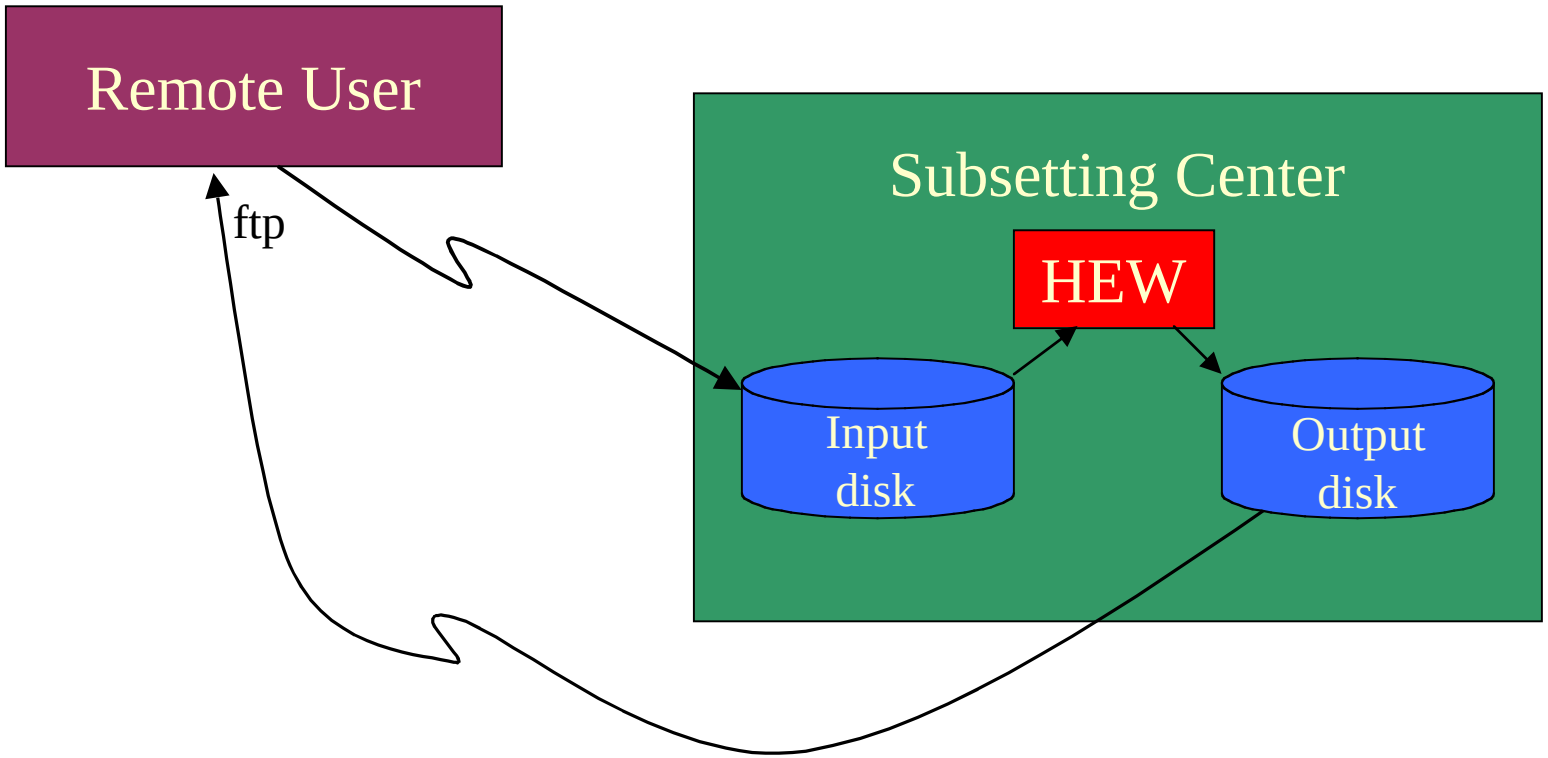
- **Subsettability "checker"**
  - **Displays content/structure of HDF-EOS files**
  - **Examines files for subsettability by HEW**
  - **Simple command-line interface**
  - **Stand-alone operation**
  - **v1.4 now available for SGI and Sun**
  - **Available at subset.org**

Subsetting Center

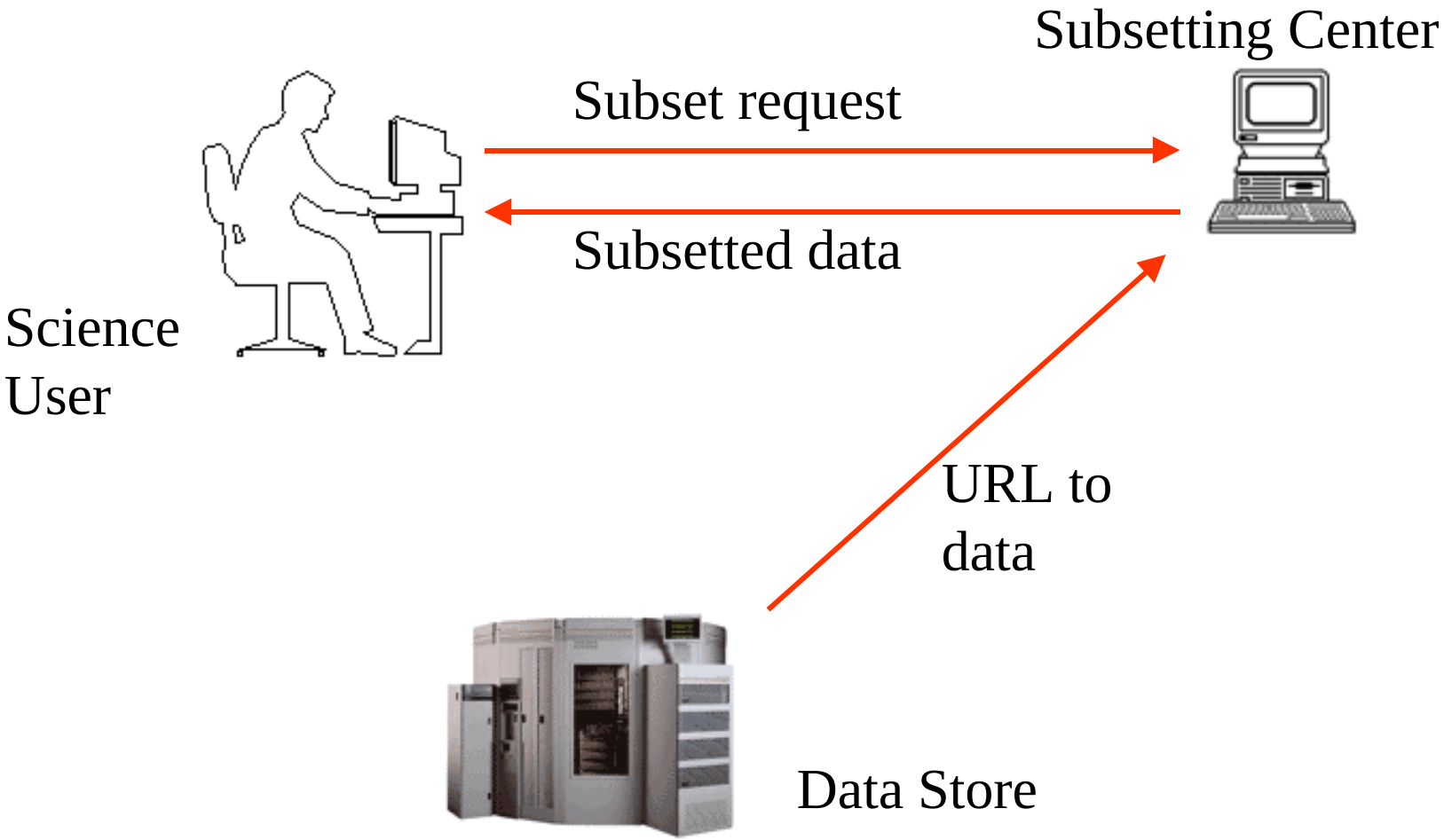
Remote user pushes file (or subsetting center could pull file) to subsetting center

Remote user interacts with HEW to perform subsetting

Remote user pulls file (or subsetting center could push file) to remote site



Subsetting as a Web Service (Planned)



HSA: HEW Subsetting Appliance

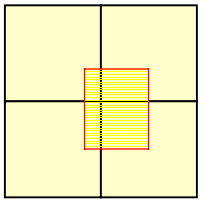
- UAH/ITSC has been working with ESDIS, ECS (EOSDIS Core System), and DAACs to incorporate subsetting appliance into the ECS/EOS Data Gateway (EDG) system.
- Provides for subsetting as part of ECS automated order filling process.
- Infrastructure written by ITSC/UAH to allow other appliances (subsetting, reprojecting, reformatting, etc.) to be added later by other data centers.

ECS integration plans

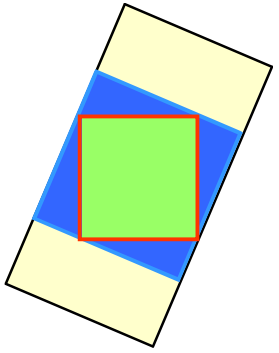
- EDG v3.4 has basic subsetting options
- Testing at NSIDC, operational in Jan ‘03
- Testing at LPDAAC (EDC)
- Testing will begin at GDAAC in ‘03
- Further enhancements as requested by DAACs

Tools developed for MODIS Scientists

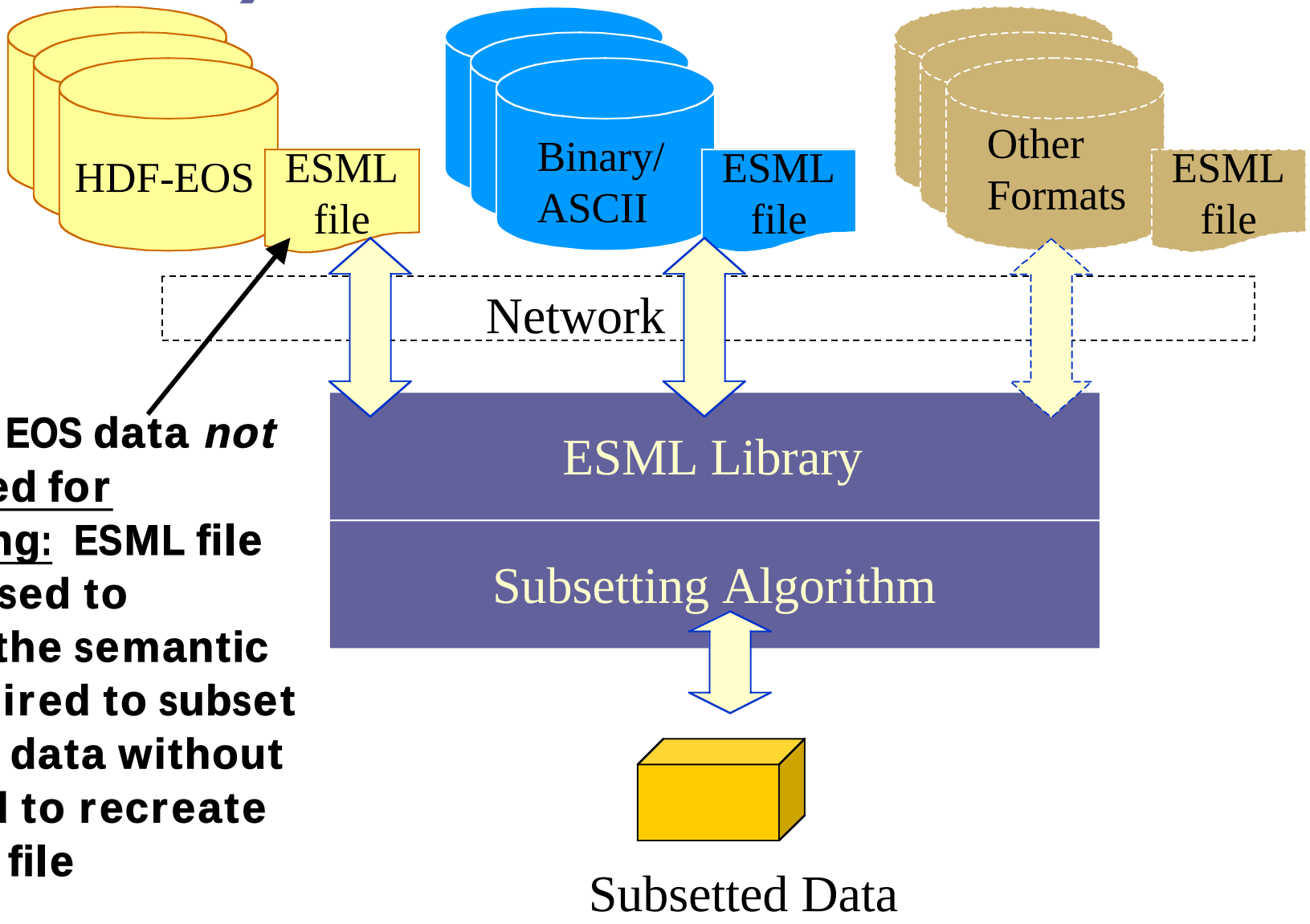
- **MODIS – Land, Quality Assessment**
  - **modland – subsetter for MODIS gridded data**
  - **stitcher – pieces together 2 or 4 contiguous MODIS tiles**



- **MODIS – Atmosphere**
  - **modair – specialized subsetter for MODIS swaths**



Subsetting Tool using ESML (in work)



For HDF-EOS data *not* formatted for subsetting: ESML file can be used to correct the semantic tag required to subset HDF-EOS data without the need to recreate the data file

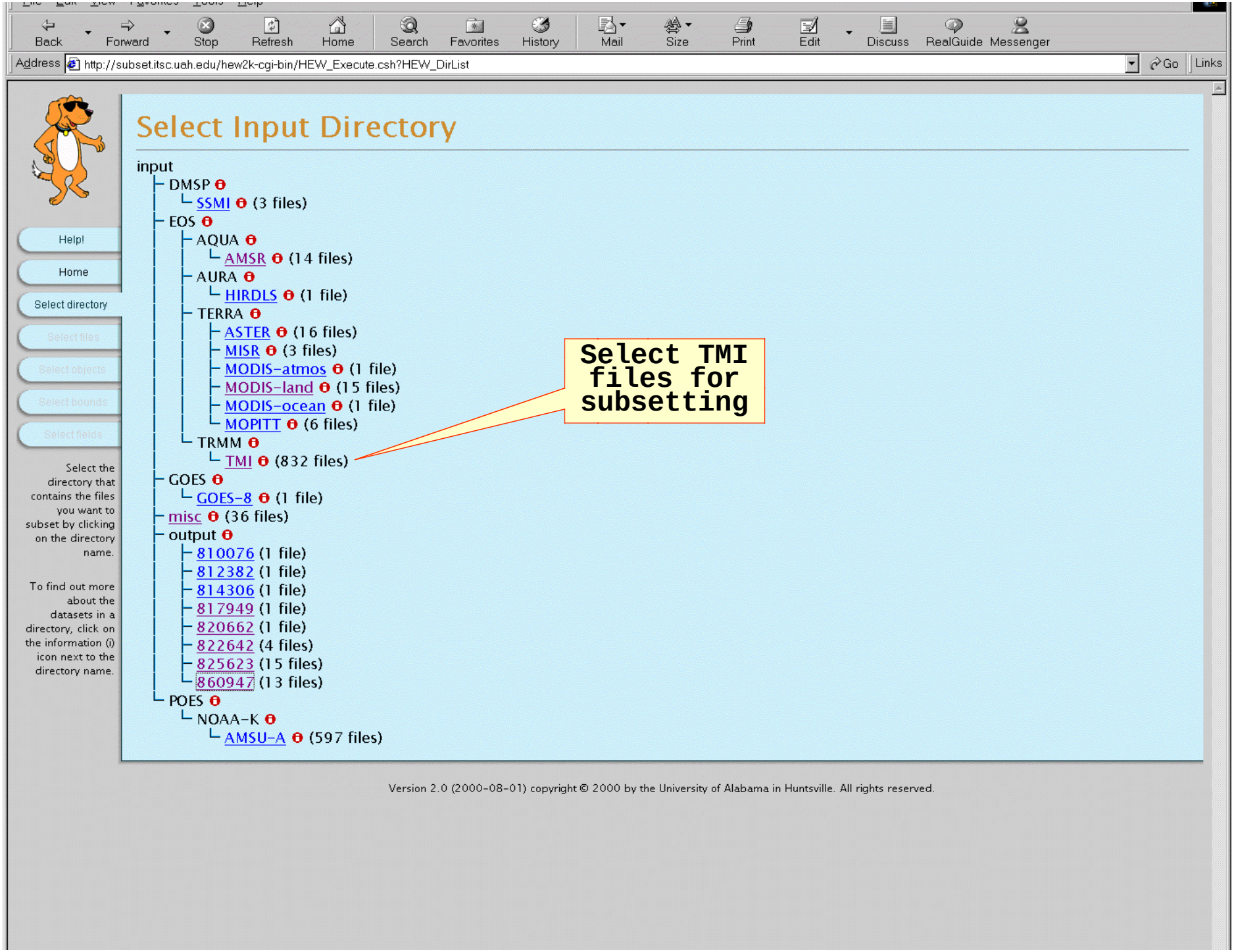
Subsetting web-site

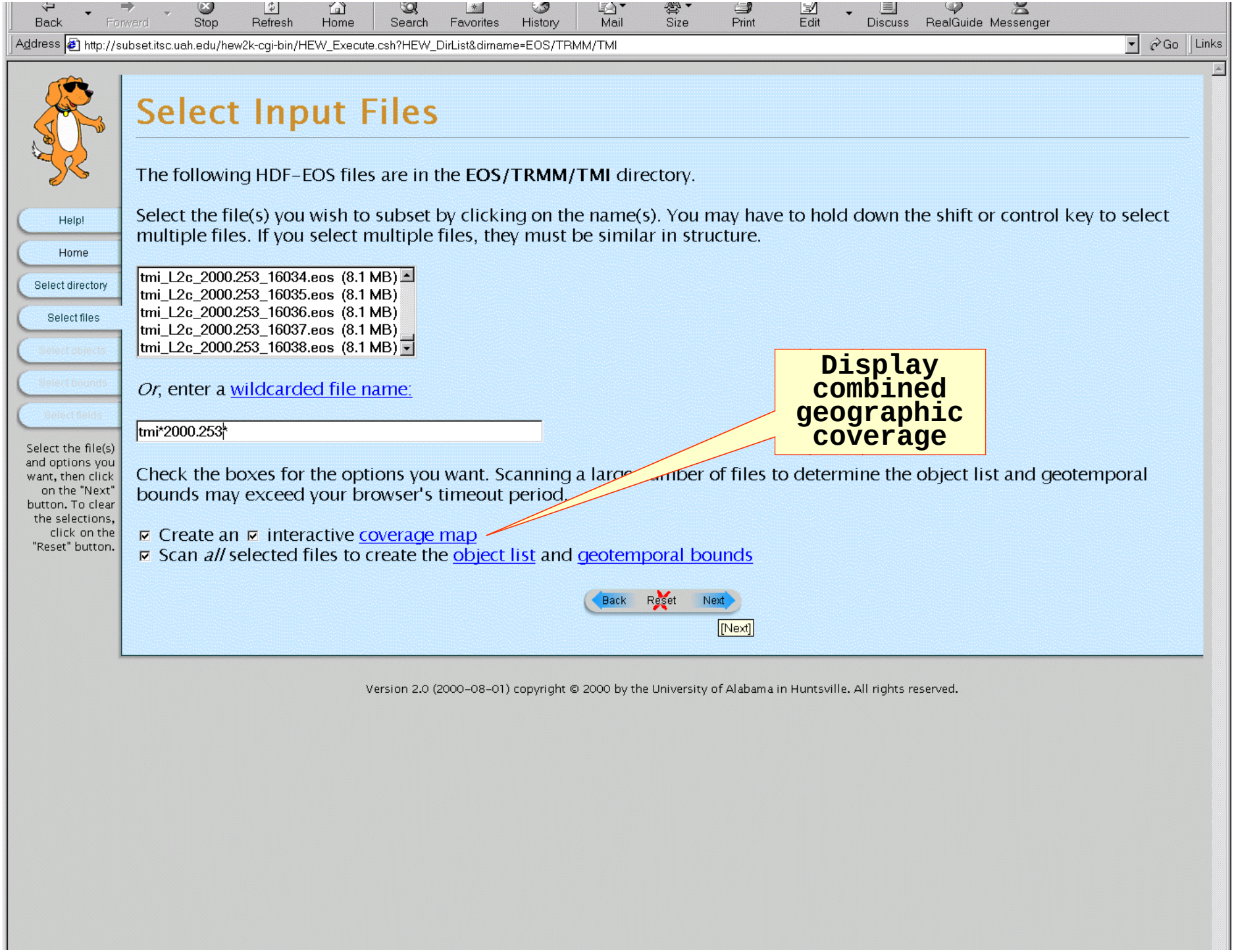
subset.org

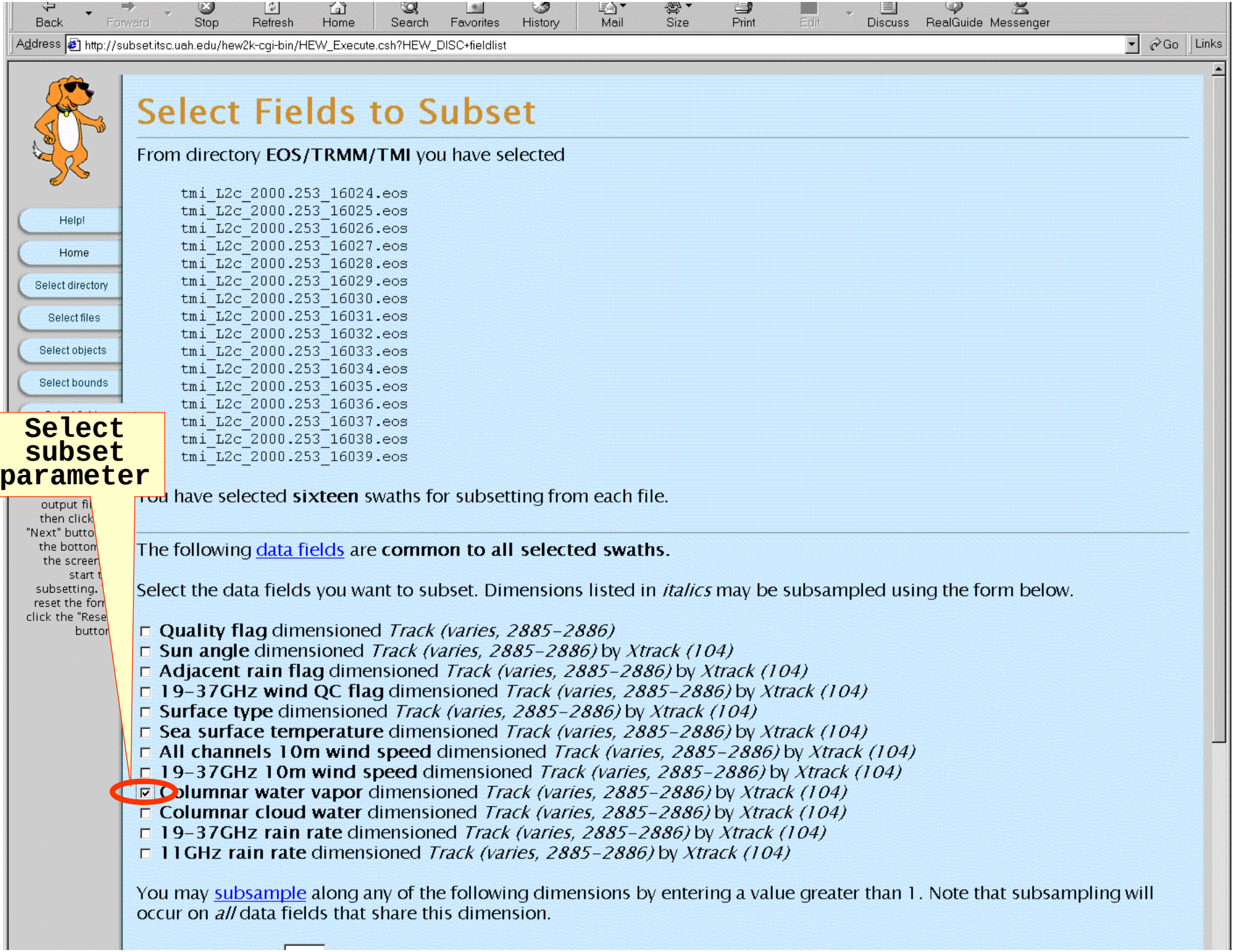
The subsetting "portal" is being created for everyone involved in subsetting

- ✓ **Advertising**
- ✓ **Forums**
- ✓ **Data**
- ✓ **Software**
- ✓ **Glossary**
- ✓ **Tutorials**
- ✓ **Links to specialized subsetters**

Subsetting Demonstration

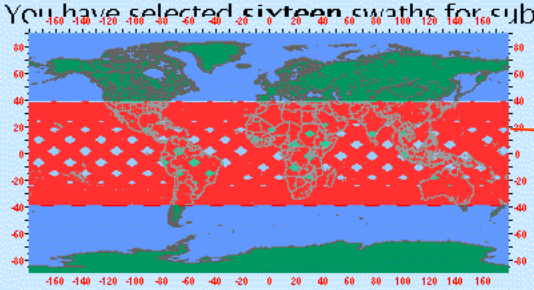

Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DirList
Go
Links
Select Input Directory
input
DMSP
SSMI (3 files)
EOS
AQUA
AMSR (14 files)
AURA
HIRDLS (1 file)
TERRA
ASTER (16 files)
MISR (3 files)
MODIS-atmos (1 file)
MODIS-land (15 files)
MODIS-ocean (1 file)
MOPITT (6 files)
TRMM
TMI (832 files)
GOES
GOES-8 (1 file)
misc (36 files)
output
810076 (1 file)
812382 (1 file)
814306 (1 file)
817949 (1 file)
820662 (1 file)
822642 (4 files)
825623 (15 files)
860947 (13 files)
POES
NOAA-K
AMSU-A (597 files)
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields
Select the directory that contains the files you want to subset by clicking on the directory name.
To find out more about the datasets in a directory, click on the information (i) icon next to the directory name.
Select TMI files for subsetting
Version 2.0 (2000-08-01) copyright © 2000 by the University of Alabama in Huntsville. All rights reserved.


Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DirList&dirname=EOS/TRMM/TMI
Go
Links
Select Input Files
The following HDF-EOS files are in the EOS/TRMM/TMI directory.
Select the file(s) you wish to subset by clicking on the name(s). You may have to hold down the shift or control key to select multiple files. If you select multiple files, they must be similar in structure.
tmi_L2c_2000.253_16034.eos (8.1 MB)
tmi_L2c_2000.253_16035.eos (8.1 MB)
tmi_L2c_2000.253_16036.eos (8.1 MB)
tmi_L2c_2000.253_16037.eos (8.1 MB)
tmi_L2c_2000.253_16038.eos (8.1 MB)
Or, enter a wildcarded file name:
tmi*2000.253*
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields
Select the file(s) and options you want, then click on the "Next" button. To clear the selections, click on the "Reset" button.
Check the boxes for the options you want. Scanning a large number of files to determine the object list and geotemporal bounds may exceed your browser's timeout period.
Create an interactive coverage map
Scan all selected files to create the object list and geotemporal bounds
Back
Reset
Next
[Next]
Display combined geographic coverage
Version 2.0 (2000-08-01) copyright © 2000 by the University of Alabama in Huntsville. All rights reserved.


Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DISC+fieldlist
Go
Links
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select subset parameter
Select Fields to Subset
From directory EOS/TRMM/TMI you have selected
tmi_L2c_2000.253_16024.eos
tmi_L2c_2000.253_16025.eos
tmi_L2c_2000.253_16026.eos
tmi_L2c_2000.253_16027.eos
tmi_L2c_2000.253_16028.eos
tmi_L2c_2000.253_16029.eos
tmi_L2c_2000.253_16030.eos
tmi_L2c_2000.253_16031.eos
tmi_L2c_2000.253_16032.eos
tmi_L2c_2000.253_16033.eos
tmi_L2c_2000.253_16034.eos
tmi_L2c_2000.253_16035.eos
tmi_L2c_2000.253_16036.eos
tmi_L2c_2000.253_16037.eos
tmi_L2c_2000.253_16038.eos
tmi_L2c_2000.253_16039.eos
have selected sixteen swaths for subsetting from each file.
The following data fields are common to all selected swaths.
Select the data fields you want to subset. Dimensions listed in italics may be subsampled using the form below.
Quality flag dimensioned Track (varies, 2885-2886)
Sun angle dimensioned Track (varies, 2885-2886) by Xtrack (104)
Adjacent rain flag dimensioned Track (varies, 2885-2886) by Xtrack (104)
19-37GHz wind QC flag dimensioned Track (varies, 2885-2886) by Xtrack (104)
Surface type dimensioned Track (varies, 2885-2886) by Xtrack (104)
Sea surface temperature dimensioned Track (varies, 2885-2886) by Xtrack (104)
All channels 10m wind speed dimensioned Track (varies, 2885-2886) by Xtrack (104)
19-37GHz 10m wind speed dimensioned Track (varies, 2885-2886) by Xtrack (104)
Columnar water vapor dimensioned Track (varies, 2885-2886) by Xtrack (104)
Columnar cloud water dimensioned Track (varies, 2885-2886) by Xtrack (104)
19-37GHz rain rate dimensioned Track (varies, 2885-2886) by Xtrack (104)
11GHz rain rate dimensioned Track (varies, 2885-2886) by Xtrack (104)
You may subsample along any of the following dimensions by entering a value greater than 1. Note that subsampling will occur on all data fields that share this dimension.

Select Geotemporal Bounds

You have selected sixteen swaths for subsetting from each file.



- Red areas or dots indicate coverage area
- Drag an edge or corner of highlight to
- Drag the middle of the highlight to move selecti
- For more precision, type in values in xes below

Combined geographic coverage of input files

Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields

Select the geotemporal bounds of your area of interest. Then click the "Next" button at the bottom of the screen. To reset the bounds to their default values, click the "Reset" button.

If you want to subset by spatial bounds, select the bounding rectangle of your area of interest:

Top
38.53
38.530

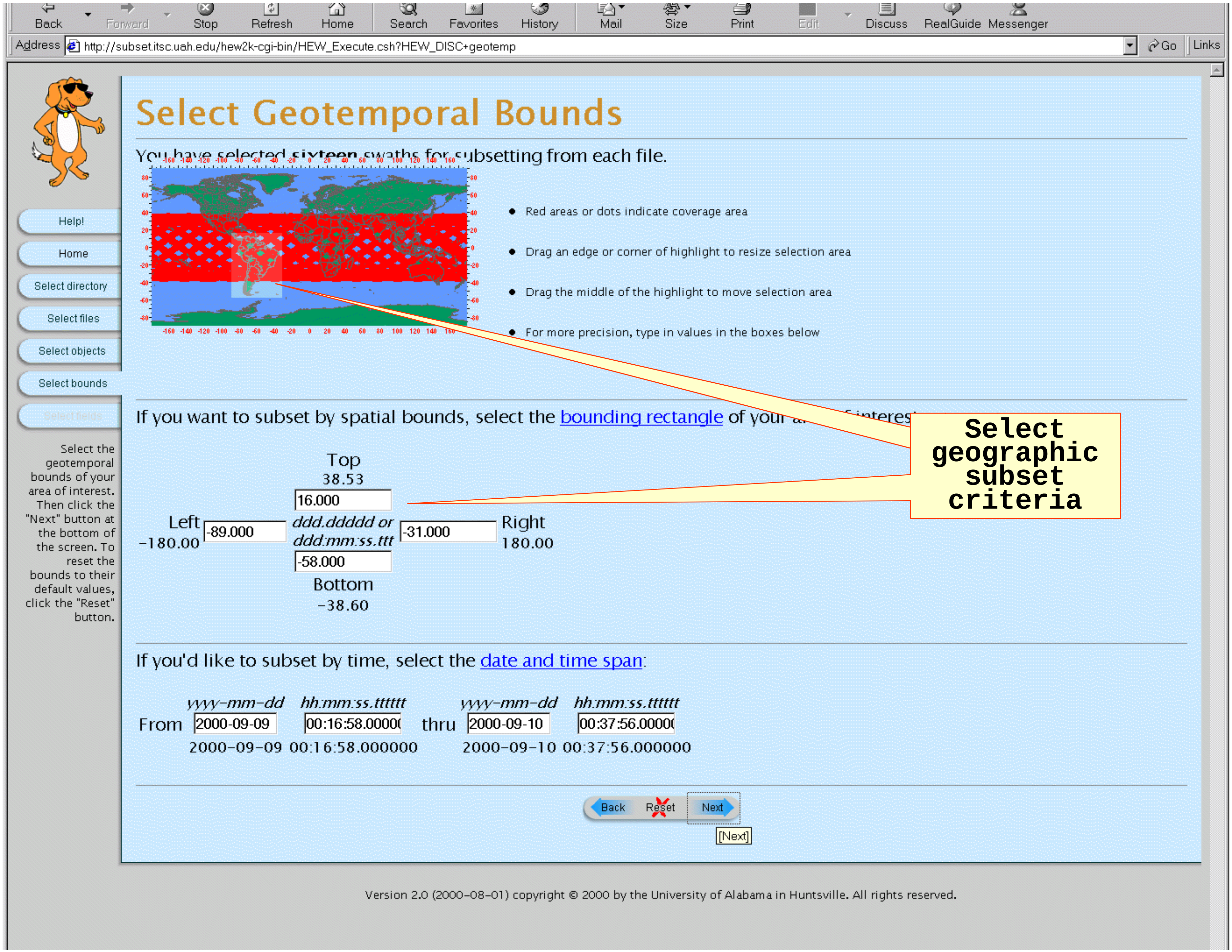
Left
-180.00
-180.000
ddd.ddddd or ddd:mm:ss.ttt
180.000
Right
180.00

-38.600
Bottom
-38.60

If you'd like to subset by time, select the date and time span:

	yyyy-mm-dd	*hh:mm:ss.tttttt*		*yyyy-mm-dd*	*hh:mm:ss.tttttt*
From	2000-09-09	00:16:58.00000	thru	2000-09-10	00:37:56.00000
	2000-09-09 00:16:58.000000			2000-09-10 00:37:56.000000	

Back Reset Next


Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DISC+geotemp
Go
Links
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select the geotemporal bounds of your area of interest. Then click the "Next" button at the bottom of the screen. To reset the bounds to their default values, click the "Reset" button.
Select Geotemporal Bounds
You have selected sixteen swaths for subsetting from each file.
Red areas or dots indicate coverage area
Drag an edge or corner of highlight to resize selection area
Drag the middle of the highlight to move selection area
For more precision, type in values in the boxes below
If you want to subset by spatial bounds, select the bounding rectangle of your a
Top
38.53
16.000
Left
-180.00
-89.000
ddd.ddddd or
ddd:mm:ss.ttt
-31.000
Right
180.00
-58.000
Bottom
-38.60
If you'd like to subset by time, select the date and time span:
yyyy-mm-dd
hh:mm:ss.tttttt
From
2000-09-09
00:16:58.00000
2000-09-09 00:16:58.000000
thru
2000-09-10
00:37:56.00000
2000-09-10 00:37:56.000000
Back
Reset
Next
[Next]
Select geographic subset criteria
Version 2.0 (2000-08-01) copyright © 2000 by the University of Alabama in Huntsville. All rights reserved.

Help!

Home

Select directory

Select files

Select objects

Select bounds

Select fields

Your subsetting request has been accepted.

Subsetting Request Submitted

Your subsetting request has been accepted. Its subsetting ID is

811280

You will need the subsetting ID to check the status of your request.

You can also check on the status of your subsetting request by using this link. If you go there now, you may be able to record the URL in your bookmarks for later use. *(Not all browsers support this function.)*

Thanks for using HEW!

Back

Home
About HEW
Sample data
Subset
Check

Select one of the tabs above to learn more about HEW or the sample datasets or to submit or check on a subsetting job.

Job Running!

Your subsetting job (811280) has been running since 15:36:01 UTC on 2000-09-15 (for about 16 seconds). Sorry, we cannot predict when it might be completed. If you wish to **cancel** the job, click on the button below.

CANCEL job 811280

Progress so far:

- 2000-09-15 15:36:03: EOS/TRMM/TMI/tmi_L2c_2000.253_16024.eos -> 811280/tmi_L2c_2000.253_16024.eos (631.3 KB)
- 2000-09-15 15:36:05: EOS/TRMM/TMI/tmi_L2c_2000.253_16025.eos -> 811280/tmi_L2c_2000.253_16025.eos (710.1 KB)
- 2000-09-15 15:36:07: EOS/TRMM/TMI/tmi_L2c_2000.253_16026.eos -> 811280/tmi_L2c_2000.253_16026.eos (715.3 KB)
- 2000-09-15 15:36:09: EOS/TRMM/TMI/tmi_L2c_2000.253_16027.eos -> 811280/tmi_L2c_2000.253_16027.eos (487.0 KB)
- 2000-09-15 15:36:11: EOS/TRMM/TMI/tmi_L2c_2000.253_16028.eos -> 811280/tmi_L2c_2000.253_16028.eos (236.3 KB)
- 2000-09-15 15:36:13: EOS/TRMM/TMI/tmi_L2c_2000.253_16029.eos -> Region or Time not found
- 2000-09-15 15:36:15: EOS/TRMM/TMI/tmi_L2c_2000.253_16030.eos -> Region or Time not found

You may click on a name to download that file to your workstation via FTP.

Note: To force your browser to download the file instead of displaying it, you may have to hold down the shift key while clicking on the file name.

Check again

This page will automatically refresh every 15 seconds.

Job Completed!

Your subsetting job (811280) **completed successfully** at 15:36:34 UTC on 2000-09-15 with the following exceptions:

- File 'EOS/TRMM/TMI/tmi_L2c_2000.253_16029.eos': Region or Time not found
- File 'EOS/TRMM/TMI/tmi_L2c_2000.253_16030.eos': Region or Time not found

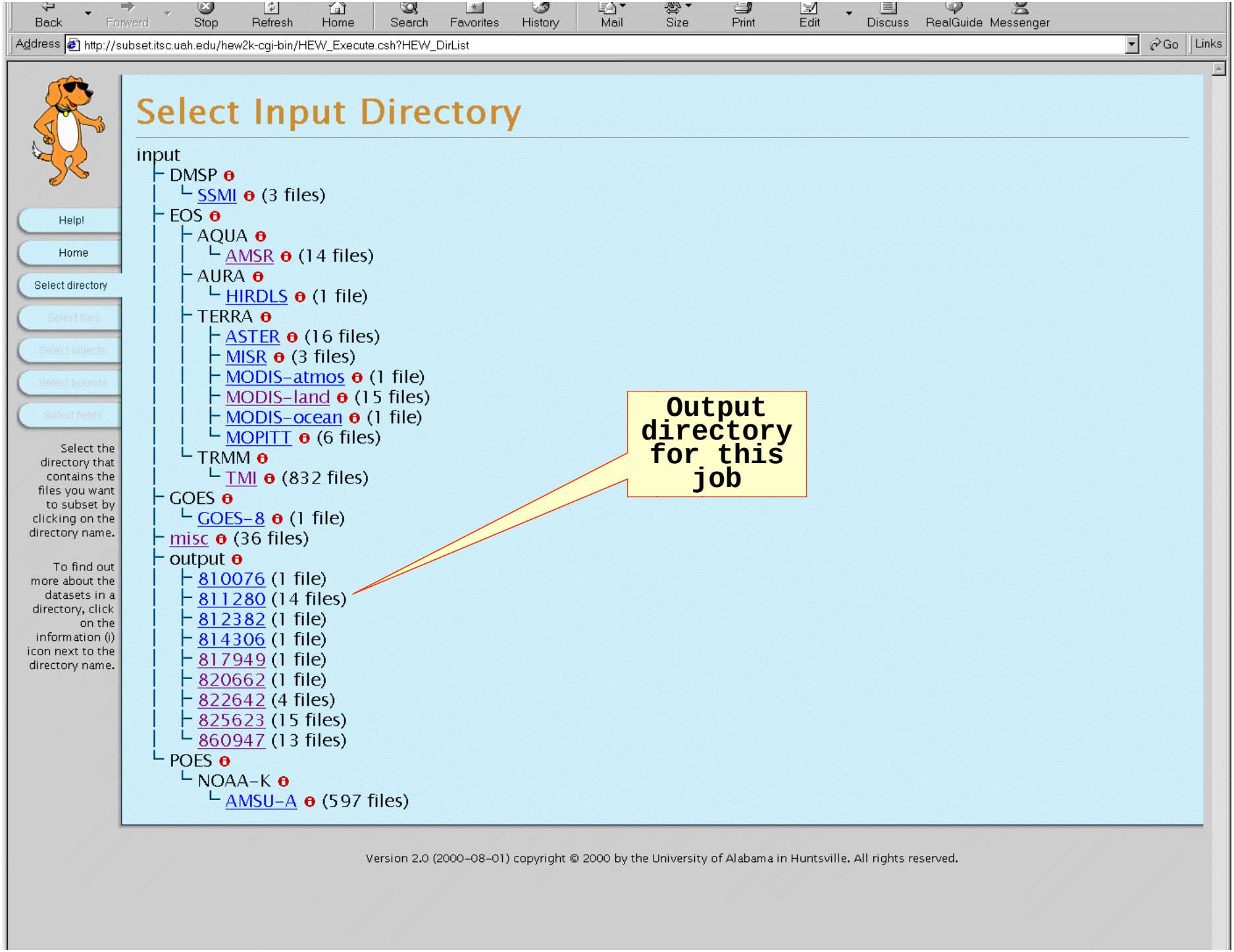
14 output files were created:

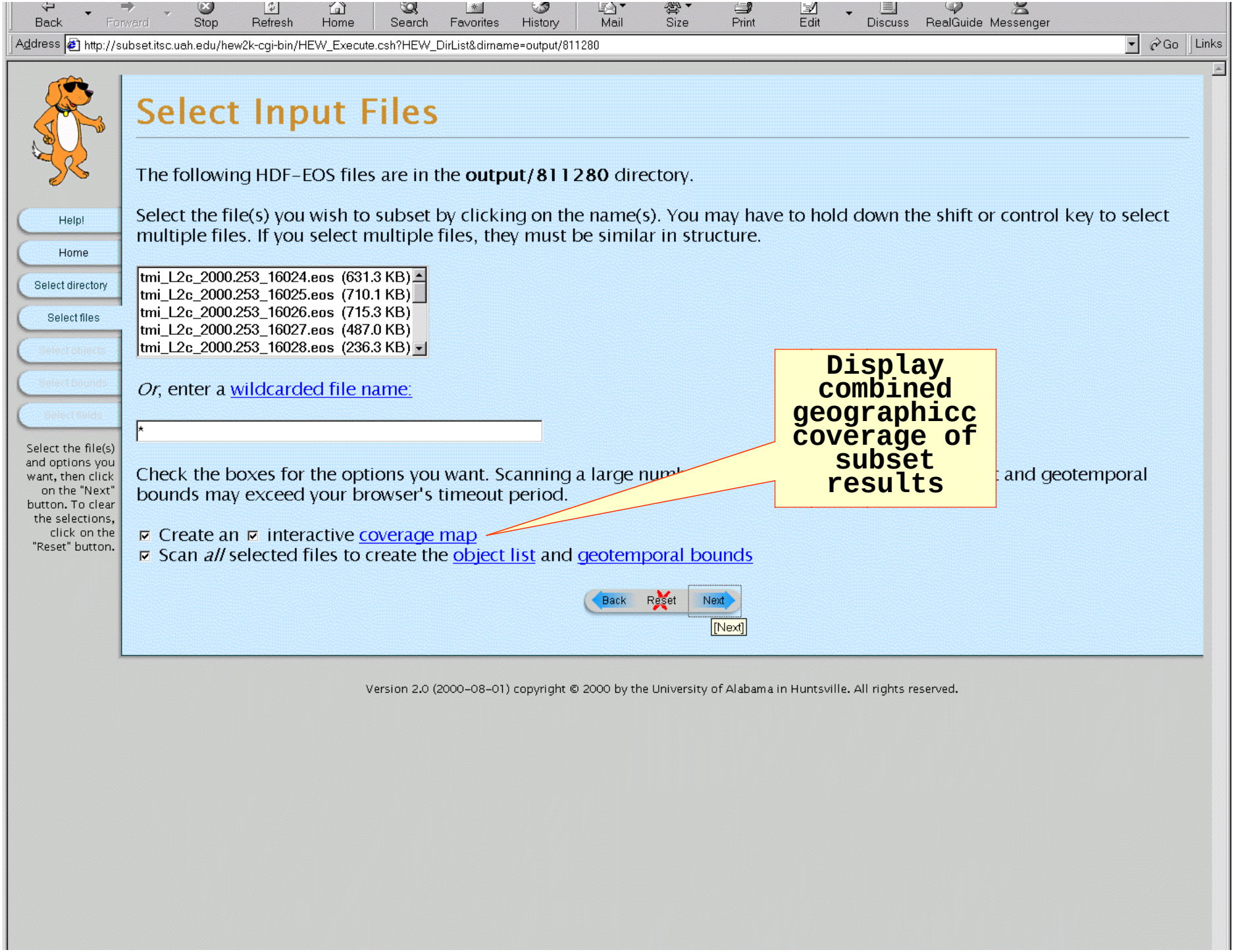
- 811280/tmi_L2c_2000.253_16024.eos (631.3 KB)
- 811280/tmi_L2c_2000.253_16025.eos (710.1 KB)
- 811280/tmi_L2c_2000.253_16026.eos (715.3 KB)
- 811280/tmi_L2c_2000.253_16027.eos (487.0 KB)
- 811280/tmi_L2c_2000.253_16028.eos (236.3 KB)
- 811280/tmi_L2c_2000.253_16031.eos (154.4 KB)
- 811280/tmi_L2c_2000.253_16032.eos (403.1 KB)
- 811280/tmi_L2c_2000.253_16033.eos (645.7 KB)
- 811280/tmi_L2c_2000.253_16034.eos (724.5 KB)
- 811280/tmi_L2c_2000.253_16035.eos (645.7 KB)
- 811280/tmi_L2c_2000.253_16036.eos (527.0 KB)
- 811280/tmi_L2c_2000.253_16037.eos (380.6 KB)
- 811280/tmi_L2c_2000.253_16038.eos (390.8 KB)
- 811280/tmi_L2c_2000.253_16039.eos (548.4 KB)

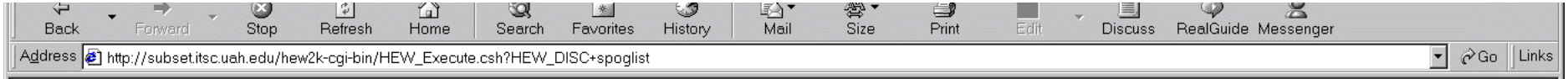
You may click on the name to download that file to your workstation via FTP.

Note: To force your browser to download the file instead of displaying it, you may have to hold down the shift key while clicking on the file name.

Do you want to submit another subsetting request?


Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DirList
Go
Links
Select Input Directory
input
DMSP
SSMI (3 files)
EOS
AQUA
AMSR (14 files)
AURA
HIRDLS (1 file)
TERRA
ASTER (16 files)
MISR (3 files)
MODIS-atmos (1 file)
MODIS-land (15 files)
MODIS-ocean (1 file)
MOPITT (6 files)
TRMM
TMI (832 files)
GOES
GOES-8 (1 file)
misc (36 files)
output
810076 (1 file)
811280 (14 files)
812382 (1 file)
814306 (1 file)
817949 (1 file)
820662 (1 file)
822642 (4 files)
825623 (15 files)
860947 (13 files)
POES
NOAA-K
AMSU-A (597 files)
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields
Select the directory that contains the files you want to subset by clicking on the directory name.
To find out more about the datasets in a directory, click on the information (i) icon next to the directory name.
Output directory for this job
Version 2.0 (2000-08-01) copyright © 2000 by the University of Alabama in Huntsville. All rights reserved.


Back
Forward
Stop
Refresh
Home
Search
Favorites
History
Mail
Size
Print
Edit
Discuss
RealGuide
Messenger
Address
http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DirList&dirname=output/811280
Go
Links
Select Input Files
The following HDF-EOS files are in the output/811280 directory.
Select the file(s) you wish to subset by clicking on the name(s). You may have to hold down the shift or control key to select multiple files. If you select multiple files, they must be similar in structure.
tmi_L2c_2000.253_16024.eos (631.3 KB)
tmi_L2c_2000.253_16025.eos (710.1 KB)
tmi_L2c_2000.253_16026.eos (715.3 KB)
tmi_L2c_2000.253_16027.eos (487.0 KB)
tmi_L2c_2000.253_16028.eos (236.3 KB)
Or, enter a wildcarded file name:
*
Check the boxes for the options you want. Scanning a large num
and geotemporal bounds may exceed your browser's timeout period.
Create an interactive coverage map
Scan all selected files to create the object list and geotemporal bounds
Back
Reset
Next
[Next]
Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields
Select the file(s) and options you want, then click on the "Next" button. To clear the selections, click on the "Reset" button.
Display combined geographicc coverage of subset results
Version 2.0 (2000-08-01) copyright © 2000 by the University of Alabama in Huntsville. All rights reserved.



Select Objects to Subset

From directory **output/811280** you have selected the following files:

```
tmi_L2c_2000.253_16024.eos
tmi_L2c_2000.253_16025.eos
tmi_L2c_2000.253_16026.eos
tmi_L2c_2000.253_16027.eos
tmi_L2c_2000.253_16028.eos
tmi_L2c_2000.253_16031.eos
tmi_L2c_2000.253_16032.eos
tmi_L2c_2000.253_16033.eos
tmi_L2c_2000.253_16034.eos
tmi_L2c_2000.253_16035.eos
tmi_L2c_2000.253_16036.eos
tmi_L2c_2000.253_16037.eos
tmi_L2c_2000.253_16038.eos
tmi_L2c_2000.253_16039.eos
```

Your file selection contains the following objects. You may subset them all at once (using the same criteria) or individually in separate operations. Please select one or more of them to subset now. You may have to hold down the shift or control key to select multiple objects.

Orbit 16028 (swath)
Orbit 16031 (swath)
Orbit 16032 (swath)
Orbit 16033 (swath)
Orbit 16034 (swath)
Orbit 16035 (swath)
Orbit 16036 (swath)
Orbit 16037 (swath)
Orbit 16038 (swath)
Orbit 16039 (swath)

Back | All | Reset | Next

Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields

Select the objects you want to subset by clicking on the object names. To select **all** objects, click the "All" button at the bottom of the screen. When you have finished making your selections, click the "Next" button.

Back Forward Stop Refresh Home Search Favorites History Mail Size Print Edit Discuss RealGuide Messenger

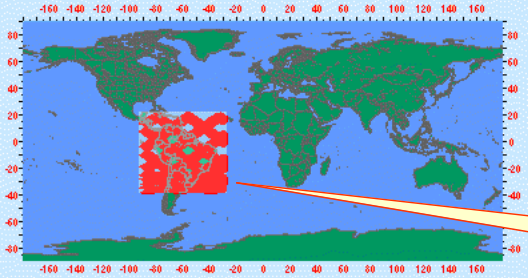
Address http://subset.itsc.uah.edu/hew2k-cgi-bin/HEW_Execute.csh?HEW_DISC+geotemp Go Links

Help!
Home
Select directory
Select files
Select objects
Select bounds
Select fields

Select the geotemporal bounds of your area of interest. Then click the "Next" button at the bottom of the screen. To reset the bounds to their default values, click the "Reset" button.

Select Geotemporal Bounds

You have selected **fourteen** swaths for subsetting from each file.



- Red areas or dots indicate coverage area
- Drag an edge or corner of highlight to resize selection area
- Drag the middle of the highlight to move selection area
- For more p… in the boxes below

Combined geographic coverage of subset results

If you want to subset by spatial bounds, select the bounding rectangle of your area of inter…

Top
21.86
21.860

Left -93.330 -93.33 | *ddd.ddddd or ddd:mm:ss.ttt* | -26.660 Right -26.66

-38.590
Bottom
-38.59

If you'd like to subset by time, select the date and time span:

	yyyy-mm-dd	*hh:mm:ss.tttttt*		*yyyy-mm-dd*	*hh:mm:ss.tttttt*
From	2000-09-09	00:17:34.00000	thru	2000-09-09	23:20:18.00000
	2000-09-09	00:17:34.000000		2000-09-09	23:20:18.000000

Back Reset Next